

A TRANSFORMATION OF THE PROBLEM OF LAGRANGE IN THE CALCULUS OF VARIATIONS*

BY

LAWRENCE M. GRAVES

By means of a simple transformation suggested by Bliss, the problem of Lagrange may be reduced to one in which the side conditions are integral equations rather than differential equations, and no derivatives enter explicitly. A multiplier rule for the transformed problem is derived below, in which the multipliers are all constants. When the inverse transformation is applied to this multiplier rule, formulas are obtained for the non-constant multipliers occurring in the ordinary form of the Lagrange multiplier rule, and it is seen that the constant multipliers obtained here may be identified with certain constants appearing in the ordinary form of the rule.

In connection with his applications of the calculus of variations to problems in economics, Roos[†] has been led to consider a generalization of the problem of Lagrange in which integro-differential equations occur among the side conditions. The transformation and analysis given below apply with equal facility to Roos' problem.

For normal arcs an analogue of the Weierstrass condition is derived for the transformed problem. It is not necessary to assume that the minimizing arc is normal on sub-arcs. For such problems as that of Roos, no generalization of the Jacobi-Mayer condition has to my knowledge yet been obtained, though several attempts have been made.

1. The transformation of the problem. We shall start with the problem suggested by Roos, in the following form: To find necessary conditions on a curve

$$y_i = y_i(x) \quad (i = 1, \dots, n; x_1 \leq x \leq x_2),$$

which minimizes an integral

$$I = \int_{x_1}^{x_2} f(x, y, y') dx$$

in a certain class of curves satisfying the integro-differential equations

* Presented to the Society, December 31, 1930; received by the editors November 21, 1932.

† *Generalized Lagrange problems in the calculus of variations*, these Transactions, vol. 30 (1928), pp. 360-384.

$$\phi_\alpha[x, y(x), y'(x), u(x)] = 0 \quad (\alpha = 1, \dots, m < n),$$

$$u_\gamma(x) = \int_{x_1}^x P_\gamma[x, s, y(s), y'(s)] ds \quad (\gamma = 1, \dots, q),$$

and the end conditions $y_i(x_1) = y_{i1}$, $y_i(x_2) = y_{i2}$. The functions f , ϕ_α , and P_γ are supposed to have continuous first partial derivatives with respect to all their arguments in a certain region R of $(2n+q+2)$ -dimensional space. The curves admitted to consideration are supposed to be of class D' , i.e., the functions $y_i(x)$ are continuous and their derivatives $y'_i(x)$ have at most a finite number of ordinary finite discontinuities. Admissible curves are also supposed to have all their elements

$$[x, s, y(s), y'(s), u(s)] \quad (x_1 \leq s \leq x \leq x_2)$$

interior to the region R . We shall suppose also that along the minimizing curve the matrix of partial derivatives $\phi_{\alpha y'_i}$ ($\alpha = 1, \dots, m$; $i = 1, \dots, n$) has rank m . For simplicity we suppose that the minimizing curve itself is of class C' .

Then as Bliss* has shown, additional functions $\phi_r(x, y')$ ($r = m+1, \dots, n$) may be adjoined, with the same continuity properties as the original functions ϕ_α , so that the functional determinant $|\phi_{i y'_j}|$ does not vanish along the minimizing curve. Hence the equations $\phi_i(x, y, y', u) = z_i$ have a unique continuous solution

$$(1) \quad y'_i = \psi_i(x, y, u, z)$$

with (x, y, y', u, z) near the values along the minimizing curve, and the functions ψ_i have continuous first partial derivatives. If equations (1) are used to eliminate the y'_i , the integral I becomes

$$I = \int_{x_1}^{x_2} g[s, y(s), u(s), z(s)] ds$$

and the side conditions become

$$y_i(x) = y_{i1} + \int_{x_1}^x \psi_i[s, y(s), u(s), z(s)] ds \quad (i = 1, \dots, n),$$

$$u_\gamma(x) = \int_{x_1}^x Q_\gamma[x, s, y(s), u(s), z(s)] ds \quad (\gamma = 1, \dots, q),$$

$$z_\alpha(x) = 0 \quad (\alpha = 1, \dots, m).$$

The end conditions are $y_i(x_2) = y_{i2}$ ($i = 1, \dots, n$).

* *The problem of Mayer with variable end points*, these Transactions, vol. 19 (1918), p. 312.

2. The multiplier rule, and an analogue of the Weierstrass condition, for the transformed problem. We shall now consider the new form of the problem on its own merits, and in order to simplify the notation in this section, we reformulate it as follows: To find necessary conditions on a "curve"

$$(2) \quad y_i = y_i(x), \quad z_r = z_r(x) \quad (i = 1, \dots, n; r = 1, \dots, \nu),$$

which minimizes an integral

$$I = \int_{x_1}^{x_2} g[s, y(s), z(s)] ds$$

in a certain class of curves satisfying the integral equations

$$(3) \quad y_i(x) = y_{i1} + \int_{x_1}^x \psi_i[x, s, y(s), z(s)] ds \quad (i = 1, \dots, n; x_1 \leq x \leq x_2)$$

and the end conditions

$$(4) \quad y_i(x_2) = y_{i2} \quad (i = 1, \dots, p \leq n).$$

The functions $g(s, y, z)$ and $\psi_i(x, s, y, z)$ are supposed to be continuous and to have continuous partial derivatives with respect to their arguments y_i and z_r in a certain region R of (x, s, y, z) space. The curves (2) admitted to consideration are supposed to have all their elements $(x, s, y(s), z(s))$ interior to R , and the functions $y_i(x)$ are supposed to be continuous, while the functions $z_r(x)$ have at most a finite number of ordinary finite discontinuities.

Under these circumstances, the equations (3) have a unique solution $y_i(x) = Y_i[z|x]$ for each set of functions $z_r(x)$ near those associated with the minimizing curve, and the functionals Y_i have differentials* $\eta_i(x) = dY_i[z; \zeta|x]$ which satisfy the equations of variation

$$(5) \quad \eta_i(x) = \int_{x_1}^x \psi_{iy_j}(x, s) \eta_j(s) ds + \int_{x_1}^x \psi_{iz_r}(x, s) \zeta_r(s) ds.$$

Here and elsewhere we abbreviate such expressions as $\psi_{iy_j}[x, s, y(s), z(s)]$ to $\psi_{iy_j}(x, s)$. When the functionals $Y_i[z]$ are substituted in the integral I , it becomes a functional $J[z]$, which is to be minimized in the class of functions $z_r(x)$ for which $Y_i[z|x_2] = y_{i2}$ ($i = 1, \dots, p$). The functional $J[z]$ also has a differential given by

$$dJ[z; \zeta] = \int_{x_1}^{x_2} \{g_{y_j}(s) \eta_j(s) + g_{z_r}(s) \zeta_r(s)\} ds,$$

* See, e.g., Graves, *Implicit functions and differential equations in general analysis*, these Transactions, vol. 29 (1927), pp. 514-552.

where the η_i are determined by equations (5). If $J[z]$ is a minimum, the usual argument shows that there exist constants $l_0, c_1, \dots, c_p$, not all zero, such that

$$(6) \quad l_0 dJ[z; \zeta] + \sum_{i=1}^p c_i dY_i[z; \zeta | x_2] = 0$$

for all functions ζ_r having only a finite number of finite discontinuities. If we set

$$(7) \quad G(s, y, z) = l_0 g(s, y, z) + \sum_{i=1}^p c_i \psi_i(x_2, s, y, z),$$

equation (6) becomes

$$(8) \quad \int_{x_1}^{x_2} \{G_{y_j}(s) \eta_j(s) + G_{z_r}(s) \zeta_r(s)\} ds = 0.$$

Let $S_{ij}(x, s)$ denote the reciprocal kernel matrix for the Volterra system (5), so that its solution is given by

$$(9) \quad \eta_i(x) = \int_{x_1}^x \psi_{iz_r}(x, s) \zeta_r(s) ds \\ - \int_{x_1}^x S_{ij}(x, t) \int_{x_1}^t \psi_{iz_r}(t, s) \zeta_r(s) ds dt.$$

By substituting (9) in (8) and making certain interchanges in the order of integration, we find

$$\int_{x_1}^{x_2} \zeta_r(x) \lambda_r(x) dx = 0$$

for all $\zeta_r(x)$, where

$$(10) \quad \lambda_r(x) = G_{z_r}(x) + \int_x^{x_2} G_{y_j}(t) \psi_{iz_r}(t, x) dt \\ - \int_x^{x_2} \int_x^u G_{y_j}(u) S_{jk}(u, t) \psi_{kz_r}(t, x) dt du.$$

Hence we have proved the

ANALOGUE OF THE LAGRANGE MULTIPLIER RULE. *If the functions $y_i(x)$, $z_r(x)$ minimize the integral I in the class of all such functions satisfying the integral equations (3) and the end conditions (4), then there exist constants $l_0, c_1, \dots, c_p$, not all zero, such that*

$$\begin{aligned}
 (11) \quad & G_{z_r}(x) + \int_x^{x_2} G_{y_j}(t) \psi_{iz_r}(t, x) dt \\
 & - \int_x^{x_2} \int_x^u G_{y_j}(u) S_{jk}(u, t) \psi_{kz_r}(t, x) dt du = 0 \\
 & (x_1 \leq x \leq x_2; r = 1, \dots, \nu),
 \end{aligned}$$

where $G(s, y, z)$ is defined by equation (7), and $S_{jk}(x, s)$ is the reciprocal kernel matrix for the system (5).

We shall say that a curve $y_i = y_i(x)$, $z_r = z_r(x)$ ($x_1 \leq x \leq x_2$) is *normal* in case there exist p sets of variations $\eta_{i\sigma}(x)$, $\zeta_{r\sigma}(x)$ ($\sigma = 1, \dots, p$), satisfying the equations of variation (5) and such that the determinant $|\eta_{i\sigma}(x_2)|$ ($i, \sigma = 1, \dots, p$) does not vanish. The usual considerations show that an arc is normal if and only if it has no set of multipliers $l_0, c_1, \dots, c_p$, with $l_0 = 0$, with which it satisfies the equations (11). For a normal minimizing arc we may always assume $l_0 = 1$, and then the remaining multipliers are uniquely determined.

ANALOGUE OF THE WEIERSTRASS CONDITION. *If the minimizing curve for our problem is normal, and if l_0 is taken equal to unity, then for every element (x, y, z) of the minimizing curve and for arbitrary numbers Z_r , the expression*

$$\begin{aligned}
 & G(x, y, Z) - G(x, y, z) \\
 & + \int_x^{x_2} \left[G_{y_j}(t) - \int_t^{x_2} G_{y_i}(s) S_{ij}(s, t) ds \right] [\psi_j(t, x, y, Z) - \psi_j(t, x, y, z)] dt
 \end{aligned}$$

cannot be negative.

This theorem may be proved by the method of the author's paper *The Weierstrass condition for the problem of Bolza in the calculus of variations*†, as follows. Let $\eta_{i\sigma}, \zeta_{r\sigma}$ ($\sigma = 1, \dots, p$) be an admissible set of functions satisfying the equations of variation (5), and such that the determinant $|\eta_{i\sigma}(x_2)| \neq 0$, where $i, \sigma = 1, \dots, p$. Let $x_1 < x_3 < x_2$, and

$$\begin{aligned}
 z_r^*(x, \beta, \epsilon) &= z_r(x) + \epsilon_r \zeta_{r\sigma}(x) \quad \text{on } x_1 \leq x \leq x_3, \quad x_3 + \beta < x \leq x_2, \\
 &= Z_r \quad \quad \quad \text{on } x_3 < x \leq x_3 + \beta.
 \end{aligned}$$

When the functions z^* are substituted for z in equations (3) these equations determine functions $y_i = y_i^*(x, \beta, \epsilon)$ defined for $x_1 \leq x \leq x_2$, (β, ϵ) near $(0, 0)$, which are continuous and have partial derivatives with respect to β and ϵ_r which are continuous except that the partial derivatives $y_{i\beta}^*$ may be discon-

† *Annals of Mathematics*, vol. 33 (1932), pp. 747-752.

tinuous in x at $x = x_3 + \beta$. Set $I(y^*, z^*) = I(\beta, \epsilon)$. We are supposing that $I(\beta, \epsilon)$ has a minimum for $\beta = \epsilon_\sigma = 0$. Then if the equations

$$(12) \quad I(\beta, \epsilon) = I(0, 0) + v, \quad y_i^*(x_2, \beta, \epsilon) = y_{i2} \quad (i = 1, \dots, p)$$

have a solution $\beta(v), \epsilon_\sigma(v)$ near $v=0$, we must have $\beta'(0) \geq 0$. By differentiating equations (12) with respect to v , we find for $\beta = \epsilon_\sigma = 0$,

$$\begin{aligned} I_\beta \beta' + I_{\epsilon_\sigma} \epsilon'_\sigma &= 1, \\ \theta_i(x_2) \beta' + \eta_{i\sigma}(x_2) \epsilon'_\sigma &= 0 \end{aligned} \quad (i = 1, \dots, p),$$

where $\theta_i(x) \equiv y_{i\beta}^*(x, 0, 0)$. Multiply the last p equations by the constants $c_1, \dots, c_p$ respectively, and add to the first. By equation (6) the result is

$$\beta' [I_\beta + \sum_{i=1}^p c_i \theta_i(x_2)] = 1.$$

Hence

$$E \equiv I_\beta + \sum_{i=1}^p c_i \theta_i(x_2) \geq 0.$$

Now the functions $\theta_i(x)$ satisfy the equations

$$\begin{aligned} \theta_i(x) &= 0 & (x_1 \leq x < x_3), \\ \theta_i(x) &= \Delta_i(x) + \int_{x_3}^x \psi_{i\nu_j}(x, s) \theta_j(s) ds & (x_3 \leq x \leq x_2), \end{aligned}$$

where $\Delta_i(x) = \psi_i(x, x_3, y_3, Z) - \psi_i(x, x_3, y_3, z_3)$, $y_3 = y(x_3)$, $z_3 = z(x_3)$. Hence by use of the reciprocal kernel S_{ij} of $\psi_{i\nu_j}$,

$$\theta_i(x) = \Delta_i(x) - \int_{x_3}^x S_{ij}(x, t) \Delta_j(t) dt.$$

By direct calculation

$$I_\beta = g(x_3, y_3, Z) - g(x_3, y_3, z_3) + \int_{x_3}^{x_2} g_{\nu_j}(s) \theta_j(s) ds.$$

Combining these results we find

$$\begin{aligned} E &= G(x_3, y_3, Z) - G(x_3, y_3, z_3) \\ &\quad + \int_{x_3}^{x_2} G_{\nu_j}(x) \left[\Delta_j(x) - \int_{x_3}^x S_{jk}(x, t) \Delta_k(t) dt \right] dx, \end{aligned}$$

which reduces by an interchange of order of integration to the expression given in the theorem.

3. Application of the inverse transformation to the new multiplier rule. Returning now to the problem of §1, we shall for simplicity consider only the case when the functions ϕ_α are independent of the u 's, that is to say, the ordinary Lagrange problem with fixed end points. Then the functions ψ_i of §2 are independent of x , and $p = n$. We shall understand that the indices used here have the following ranges: $i, j, k, l = 1, \dots, n$; $\alpha = 1, \dots, m$; $r = m+1, \dots, n$. From the definitions of the functions ψ_i, S_{ij} , and G , we obtain the following relations:

$$(13) \quad \psi_{kz} \phi_{iy'_i} = \delta_{kj},$$

$$(14) \quad \psi_{kz} \phi_{jy_i} = -\psi_{ky_i},$$

$$(15) \quad \psi_{iy_l}(x) - \int_x^v S_{jk}(v, t) dt \psi_{ky_l}(x) = -S_{jl}(v, x),$$

$$(16) \quad G_{zi} = (l_0 f_{y'_i} + c_i) \psi_{iz_i},$$

$$(17) \quad G_{y_j} = l_0 f_{y_j} + (l_0 f_{y'_k} + c_k) \psi_{ky_j}.$$

The analogue of the Euler-Lagrange equations may be written

$$(18) \quad G_{zi}(x) + \psi_{kz}(x) \int_x^{x_1} \left[G_{y_k}(t) - G_{y_j}(t) \int_x^t S_{jk}(t, v) dv \right] dt = \lambda_i(x),$$

$$\lambda_r(x) = 0.$$

If we multiply equations (18) by $\phi_{iy'_i}$ and add, use equations (13) and (16), and interchange the order of integration in the double integral, we find

$$(19) \quad l_0 f_{y'_i} + c_i + \int_x^{x_1} G_{y_l}(t) dt - \int_x^{x_1} \int_t^{x_1} G_{y_j}(v) S_{jl}(v, t) dv dt = \lambda_\alpha \phi_{\alpha y'_i}.$$

Also if we multiply equations (18) by ϕ_{iy_l} and add, we find with the help of equations (14), (15) and (16),

$$(20) \quad \int_x^{x_1} G_{y_j}(v) S_{jl}(v, x) dv = (l_0 f_{y'_i} + c_i) \psi_{iy_l} + \lambda_\alpha \phi_{\alpha y_l}.$$

Combining equations (20) and (17) with (19) we find

$$l_0 f_{y'_i} + c_i + \int_x^{x_1} (l_0 f_{y_l} - \lambda_\alpha \phi_{\alpha y_l}) dt = \lambda_\alpha \phi_{\alpha y'_i}$$

which may be written in the familiar form

$$(21) \quad F_{y_i'} = \int_{x_2}^x F_{y_i} dx - c_i$$

by setting $F = l_0 f - \lambda_\alpha \phi_\alpha$.

If we apply the inverse transformation in the more general problem considered in §1, we find in place of equations (21),

$$\begin{aligned} \dot{F}_{y_i'}(x) = & \int_{x_2}^x \left[F_{y_i}(s) + \int_s^{x_1} F_{u_\gamma}(t) P_{\gamma y_i}(t, s) dt \right. \\ & \left. + F_{u_\gamma}(s) P_{\gamma y_i'}(s, x) \right] ds - c_i. \end{aligned}$$

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.